

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/816,467

DATE: 07/17/2001

TIME: 09:40:58

Input Set : A:\34950174.app

Output Set: N:\CRF3\07172001\I816467.raw

ENTERED

3 <110> APPLICANT: COEN, LAURENT
4 PINZOLAS, ROSARIO OSTA
5 BRULET, PHILIPPE
7 <120> TITLE OF INVENTION: HYBRID PROTEINS THAT MIGRATE RETROGRADELY AND
8 TRANSYNAPTICALLY INTO THE CNS
10 <130> FILE REFERENCE: 03495.0174-01000
12 <140> CURRENT APPLICATION NUMBER: 09/816,467
13 <141> CURRENT FILING DATE: 2001-03-26
15 <150> PRIOR APPLICATION NUMBER: 60/055,615
16 <151> PRIOR FILING DATE: 1997-08-14
18 <150> PRIOR APPLICATION NUMBER: 60/065,236
19 <151> PRIOR FILING DATE: 1997-11-13
21 <160> NUMBER OF SEQ ID NOS: 19
23 <170> SOFTWARE: PatentIn Ver. 2.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 1600
27 <212> TYPE: DNA
28 <213> ORGANISM: Clostridium tetani
30 <220> FEATURE:
31 <221> NAME/KEY: CDS
32 <222> LOCATION: (88)..(1476)
34 <400> SEQUENCE: 1
35 ggaaacagct atgaccatga ttacgccaag ctcgaaatta accctcacta aaggggaacaa 60
37 aagctggagc tcggtaccgc ggccacc atg gtt ttt tca aca cca att cca ttt 114
38 Met Val Phe Ser Thr Pro Ile Pro Phe
39 1 5
41 tct tat tct aaa aat ctg gat tgt tgg gtt gat aat gaa gaa gat ata 162
42 Ser Tyr Ser Lys Asn Leu Asp Cys Trp Val Asp Asn Glu Glu Asp Ile
43 10 15 20 25
45 gat gtt ata tta aaa aag agt aca att tta aat tta gat att aat aat 210
46 Asp Val Ile Leu Lys Lys Ser Thr Ile Leu Asn Leu Asp Ile Asn Asn
47 30 35 40
49 gat att ata tca gat ata tct ggg ttt aat tca tct gta ata aca tat 258
50 Asp Ile Ile Ser Asp Ile Ser Gly Phe Asn Ser Ser Val Ile Thr Tyr
51 45 50 55
53 cca gat gct caa ttg gtg ccc gga ata aat ggc aaa gca ata cat tta 306
54 Pro Asp Ala Gln Leu Val Pro Gly Ile Asn Gly Lys Ala Ile His Leu
55 60 65 70
57 gta aac aat gaa tct tct gaa gtt ata gtg cat aaa gct atg gat att 354
58 Val Asn Asn Glu Ser Ser Glu Val Ile Val His Lys Ala Met Asp Ile
59 75 80 85
61 gaa tat aat gat atg ttt aat aat ttt acc gtt agc ttt tgg ttg agg 402
62 Glu Tyr Asn Asp Met Phe Asn Asn Phe Thr Val Ser Phe Trp Leu Arg
63 90 95 100 105
65 gtt cct aaa gta tct gct agt cat tta gaa caa tat ggc aca aat gag 450
66 Val Pro Lys Val Ser Ala Ser His Leu Glu Gln Tyr Gly Thr Asn Glu
67 110 115 120

RAW SEQUENCE LISTING

DATE: 07/17/2001

PATENT APPLICATION: US/09/616,467

TIME: 09:40:58

Input Set : A:\34950174.app

Output Set: N:\CRF3\07172001\I816467.raw

```
69 tat tca ata att agc tct atg aaa aaa cat agt cta tca ata gga tct 498
70 Tyr Ser Ile Ile Ser Ser Met Lys Lys His Ser Leu Ser Ile Gly Ser
71 125 130 135
73 ggt tgg agt gta tca ctt aaa ggt aat aac tta ata tgg act tta aaa 546
74 Gly Trp Ser Val Ser Leu Lys Gly Asn Asn Leu Ile Trp Thr Leu Lys
75 140 145 150
77 gat tcc gcg gga gaa gtt aga caa ata act ttt agg gat tta cct gat 594
78 Asp Ser Ala Gly Glu Val Arg Gln Ile Thr Phe Arg Asp Leu Pro Asp
79 155 160 165
81 aaa ttt aat gct tat tta gca aat aaa tgg gtt ttt ata act att act 642
82 Lys Phe Asn Ala Tyr Leu Ala Asn Lys Trp Val Phe Ile Thr Ile Thr
83 170 175 180 185
85 aat gat aga tta tct tct gct aat ttg tat ata aat gga gta ctt atg 690
86 Asn Asp Arg Leu Ser Ser Ala Asn Leu Tyr Ile Asn Gly Val Leu Met
87 190 195 200
89 gga agt gca gaa att act ggt tta gga gct att aga gag gat aat aat 738
90 Gly Ser Ala Glu Ile Thr Gly Leu Gly Ala Ile Arg Glu Asp Asn Asn
91 205 210 215
93 ata aca tta aaa cta gat aga tgt aat aat aat aat caa tac gtt tct 786
94 Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Asn Asn Gln Tyr Val Ser
95 220 225 230
97 att gat aaa ttt agg ata ttt tgc aaa gca tta aat cca aaa gag att 834
98 Ile Asp Lys Phe Arg Ile Phe Cys Lys Ala Leu Asn Pro Lys Glu Ile
99 235 240 245
101 gaa aaa tta tac aca agt tat tta tct ata acc ttt tta aga gac ttc 882
102 Glu Lys Leu Tyr Thr Ser Tyr Leu Ser Ile Thr Phe Leu Arg Asp Phe
103 250 255 260 265
105 tgg gga aac cct tta cga tat gat aca gaa tat tat tta ata cca gta 930
106 Trp Gly Asn Pro Leu Arg Tyr Asp Thr Glu Tyr Tyr Leu Ile Pro Val
107 270 275 280
109 gct tct agt tct aaa gat gtt caa ttg aaa aat ata aca gat tat atg 978
110 Ala Ser Ser Ser Lys Asp Val Gln Leu Lys Asn Ile Thr Asp Tyr Met
111 285 290 295
113 tat ttg aca aat gcg cca tcg tat act aac gga aaa ttg aat ata tat 1026
114 Tyr Leu Thr Asn Ala Pro Ser Tyr Thr Asn Gly Lys Leu Asn Ile Tyr
115 300 305 310
117 tat aga agg tta tat aat gga cta aaa ttt att ata aaa aga tat aca 1074
118 Tyr Arg Arg Leu Tyr Asn Gly Leu Lys Phe Ile Ile Lys Arg Tyr Thr
119 315 320 325
121 cct aat aat gaa ata gat tct ttt gtt aaa tca ggt gat ttt att aaa 1122
122 Pro Asn Asn Glu Ile Asp Ser Phe Val Lys Ser Gly Asp Phe Ile Lys
123 330 335 340 345
125 tta tat gta tca tat aac aat aat gag cac att gta ggt tat ccg aaa 1170
126 Leu Tyr Val Ser Tyr Asn Asn Asn Glu His Ile Val Gly Tyr Pro Lys
127 350 355 360
129 gat gga aat gcc ttt aat aat ctt gat aga att cta aga gta ggt tat 1218
130 Asp Gly Asn Ala Phe Asn Asn Leu Asp Arg Ile Leu Arg Val Gly Tyr
131 365 370 375
133 aat gcc cca ggt atc cct ctt tat aaa aaa atg gaa gca gta aaa ttg 1266
```

RAW SEQUENCE LISTING

DATE: 07/17/2001

PATENT APPLICATION: US/09/816,467

TIME: 09:40:56

Input Set : A:\34950174.app

Output Set: N:\CRF3\07172001\I816467.raw

```
134 Asn Ala Pro Gly Ile Pro Leu Tyr Lys Lys Met Glu Ala Val Lys Leu
135          380          385          390
137 cgt gat tta aaa acc tat tct gta caa ctt aaa tta tat gat gat aaa 1314
138 Arg Asp Leu Lys Thr Tyr Ser Val Gln Leu Lys Leu Tyr Asp Asp Lys
139          395          400          405
141 aat gca tct tta gga cta gta ggt acc cat aat ggt caa ata ggc aac 1362
142 Asn Ala Ser Leu Gly Leu Val Gly Thr His Asn Gly Gln Ile Gly Asn
143 410          415          420          425
145 gat cca aat agg gat ata tta att gca agc aac tgg tac ttt aat cat 1410
146 Asp Pro Asn Arg Asp Ile Leu Ile Ala Ser Asn Trp Tyr Phe Asn His
147          430          435          440
149 tta aaa gat aaa att tta gga tgt gat tgg tac ttt gta cct aca gat 1458
150 Leu Lys Asp Lys Ile Leu Gly Cys Asp Trp Tyr Phe Val Pro Thr Asp
151          445          450          455
153 gag gga tgg aca aat gat taaacagatt gatatgttca tgacatatgc 1506
154 Glu Gly Trp Thr Asn Asp
155          460
157 ccgggatcct ctagagtcga cctcgagggg gggcccggtg cccaattcgc cctatagtga 1566
159 gtcgtattac aattcactgg ccgtcgtttt acaa 1600
162 <210> SEQ ID NO: 2
163 <211> LENGTH: 463
164 <212> TYPE: PRT
165 <213> ORGANISM: Clostridium tetani
167 <400> SEQUENCE: 2
168 Met Val Phe Ser Thr Pro Ile Pro Phe Ser Tyr Ser Lys Asn Leu Asp
169 1 5 10 15
171 Cys Trp Val Asp Asn Glu Glu Asp Ile Asp Val Ile Leu Lys Lys Ser
172 20 25 30
174 Thr Ile Leu Asn Leu Asp Ile Asn Asn Asp Ile Ile Ser Asp Ile Ser
175 35 40 45
177 Gly Phe Asn Ser Ser Val Ile Thr Tyr Pro Asp Ala Gln Leu Val Pro
178 50 55 60
180 Gly Ile Asn Gly Lys Ala Ile His Leu Val Asn Asn Glu Ser Ser Glu
181 65 70 75 80
183 Val Ile Val His Lys Ala Met Asp Ile Glu Tyr Asn Asp Met Phe Asn
184 85 90 95
186 Asn Phe Thr Val Ser Phe Trp Leu Arg Val Pro Lys Val Ser Ala Ser
187 100 105 110
189 His Leu Glu Gln Tyr Gly Thr Asn Glu Tyr Ser Ile Ile Ser Ser Met
190 115 120 125
192 Lys Lys His Ser Leu Ser Ile Gly Ser Gly Trp Ser Val Ser Leu Lys
193 130 135 140
195 Gly Asn Asn Leu Ile Trp Thr Leu Lys Asp Ser Ala Gly Glu Val Arg
196 145 150 155 160
198 Gln Ile Thr Phe Arg Asp Leu Pro Asp Lys Phe Asn Ala Tyr Leu Ala
199 165 170 175
201 Asn Lys Trp Val Phe Ile Thr Ile Thr Asn Asp Arg Leu Ser Ser Ala
202 180 185 190
204 Asn Leu Tyr Ile Asn Gly Val Leu Met Gly Ser Ala Glu Ile Thr Gly
```

RAW SEQUENCE LISTING

DATE: 07/17/2001

PATENT APPLICATION: US/09/816,467

TIME: 09:40:58

Input Set : A:\34950174.app

Output Set: N:\CRF3\07172001\I816467.raw

```
205          195          200          205
207 Leu Gly Ala Ile Arg Glu Asp Asn Asn Ile Thr Leu Lys Leu Asp Arg
208          210          215          220
210 Cys Asn Asn Asn Asn Gln Tyr Val Ser Ile Asp Lys Phe Arg Ile Phe
211 225          230          235          240
213 Cys Lys Ala Leu Asn Pro Lys Glu Ile Glu Lys Leu Tyr Thr Ser Tyr
214          245          250          255
216 Leu Ser Ile Thr Phe Leu Arg Asp Phe Trp Gly Asn Pro Leu Arg Tyr
217          260          265          270
219 Asp Thr Glu Tyr Tyr Leu Ile Pro Val Ala Ser Ser Ser Lys Asp Val
220          275          280          285
222 Gln Leu Lys Asn Ile Thr Asp Tyr Met Tyr Leu Thr Asn Ala Pro Ser
223          290          295          300
225 Tyr Thr Asn Gly Lys Leu Asn Ile Tyr Tyr Arg Arg Leu Tyr Asn Gly
226 305          310          315          320
228 Leu Lys Phe Ile Ile Lys Arg Tyr Thr Pro Asn Asn Glu Ile Asp Ser
229          325          330          335
231 Phe Val Lys Ser Gly Asp Phe Ile Lys Leu Tyr Val Ser Tyr Asn Asn
232          340          345          350
234 Asn Glu His Ile Val Gly Tyr Pro Lys Asp Gly Asn Ala Phe Asn Asn
235          355          360          365
237 Leu Asp Arg Ile Leu Arg Val Gly Tyr Asn Ala Pro Gly Ile Pro Leu
238          370          375          380
240 Tyr Lys Lys Met Glu Ala Val Lys Leu Arg Asp Leu Lys Thr Tyr Ser
241 385          390          395          400
243 Val Gln Leu Lys Leu Tyr Asp Asp Lys Asn Ala Ser Leu Gly Leu Val
244          405          410          415
246 Gly Thr His Asn Gly Gln Ile Gly Asn Asp Pro Asn Arg Asp Ile Leu
247          420          425          430
249 Ile Ala Ser Asn Trp Tyr Phe Asn His Leu Lys Asp Lys Ile Leu Gly
250          435          440          445
252 Cys Asp Trp Tyr Phe Val Pro Thr Asp Glu Gly Trp Thr Asn Asp
253          450          455          460
257 <210> SEQ ID NO: 3
258 <211> LENGTH: 1392
259 <212> TYPE: DNA
260 <213> ORGANISM: Clostridium tetani
262 <400> SEQUENCE: 3
263 atggtttttt caacaccaat tccattttct tattctaaaa atctggattg ttgggttgat 60
264 aatgaagaag atatagatgt tatattaaaa aagagtacaa ttttaaattt agatattaat 120
265 aatgatatta tatcagatat atctgggttt aattcatctg taataacata tccagatgct 180
266 caattggtgc ccggaataaa tggcaaagca atacatttag taaacaatga atcttctgaa 240
267 gttatagtgc ataaagctat ggatattgaa tataatgata tgtttaataa ttttaccgtt 300
268 agcttttggg tgagggttcc taaagtatct gctagtcatt tagaacaata tggcacaaat 360
269 gagtattcaa taattagctc tatgaaaaaa catagtctat caataggatc tggttggagt 420
270 gtatcactta aaggtaataa cttaatatgg actttaaaag attccgcggg agaagttaga 480
271 caaataactt ttagggattt acctgataaa tttaatgctt atttagcaaa taaatgggtt 540
272 ttataacta ttactaatga tagattatct tctgctaatt tgtatataaa tggagtactt 600
273 atgggaagtg cagaaattac tggtttagga gctattagag aggataataa tataacatta 660
```

RAW SEQUENCE LISTING

DATE: 07/17/2001

PATENT APPLICATION: US/09/816,467

TIME: 09:40:58

Input Set : A:\34950174.app

Output Set: N:\CRF3\07172001\I816467.raw

```
274 aaactagata gatgtaataa taataatcaa tacgtttcta ttgataaatt taggatattt 720
275 tgcaaagcat taaatccaaa agagattgaa aaattataca caagttattt atctataacc 780
276 tttttaagag acttctgggg aaacccttta cgatatgata cagaatatta tttataacca 840
277 gtagcttcta gttctaaaga tgttcaattg aaaaatataa cagattatat gtatttgaca 900
278 aatgcgccat cgtatactaa cggaaaattg aatatatatt atagaagggt atataatgga 960
279 ctaaaattta ttataaaaag atatacacct aataatgaaa tagattcttt tgttaaatca 1020
280 ggtgatttta ttaaattata tgtatcatat aacaataatg agcacattgt aggttatccg 1080
281 aaagatggaa atgcctttaa taatcttgat agaattctaa gagtaggtta taatgcccc 1140
282 ggtatccctc tttataaaaa aatggaagca gtaaaattgc gtgattttaa aacctattct 1200
283 gtacaactta aattatatga tgataaaaat gcactcttag gactagtagg taccataat 1260
284 ggtcaaatag gcaacgatoc aaatagggat atattaattg caagcaactg gtactttaat 1320
285 catttaaaag ataaaatttt aggatgtgat tggtaacttg tacctacaga tgagggatgg 1380
286 acaaatgatt aa 1392
289 <210> SEQ ID NO: 4
290 <211> LENGTH: 49
291 <212> TYPE: DNA
292 <213> ORGANISM: Artificial Sequence
294 <220> FEATURE:
295 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
297 <400> SEQUENCE: 4
298 cccccgggc caccatgggt ttttcaacac caattccatt ttcttattc 49
301 <210> SEQ ID NO: 5
302 <211> LENGTH: 18
303 <212> TYPE: DNA
304 <213> ORGANISM: Artificial Sequence
306 <220> FEATURE:
307 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
309 <400> SEQUENCE: 5
310 ctaaaccagt aatttctg 18
313 <210> SEQ ID NO: 6
314 <211> LENGTH: 25
315 <212> TYPE: DNA
316 <213> ORGANISM: Artificial Sequence
318 <220> FEATURE:
319 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
321 <400> SEQUENCE: 6
322 aattatggac tttaaaagat tccgc 25
325 <210> SEQ ID NO: 7
326 <211> LENGTH: 24
327 <212> TYPE: DNA
328 <213> ORGANISM: Artificial Sequence
330 <220> FEATURE:
331 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
333 <400> SEQUENCE: 7
334 ggcattataa cctactotta gaat 24
337 <210> SEQ ID NO: 8
338 <211> LENGTH: 27
339 <212> TYPE: DNA
340 <213> ORGANISM: Artificial Sequence
```

VERIFICATION SUMMARY

DATE: 07/17/2001

PATENT APPLICATION: US/09/816,467

TIME: 09:40:59

Input Set : A:\34950174.app

Output Set: N:\CRF3\07172001\I816467.raw